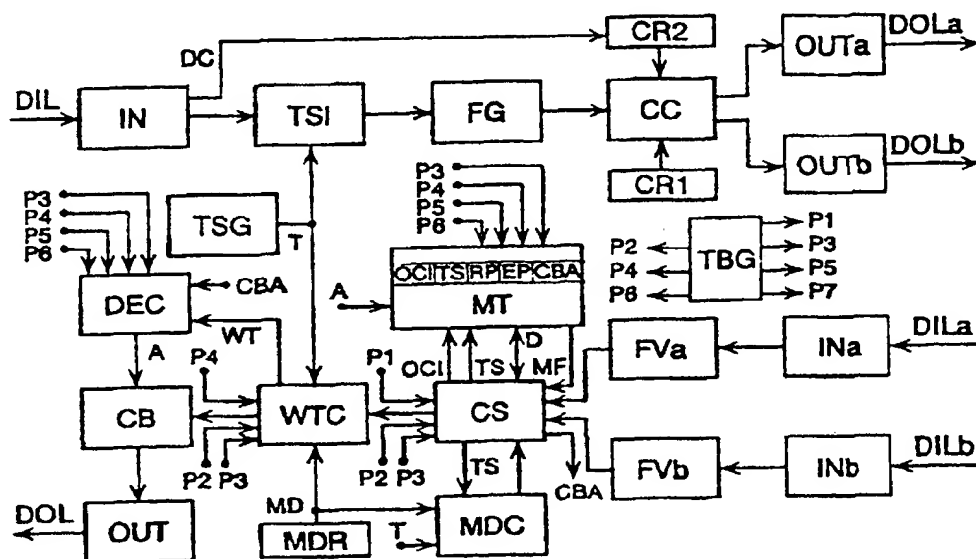




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>5</sup> : <b>H04L 12/56, H04Q 11/04</b>		<b>A1</b>	(11) International Publication Number: <b>WO 94/27391</b>
			(43) International Publication Date: 24 November 1994 (24.11.94)
(21) International Application Number: <b>PCT/EP94/00915</b>		(81) Designated States: US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 22 March 1994 (22.03.94)		<b>Published</b> With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(30) Priority Data: MI93A001001 17 May 1993 (17.05.93) IT			
(71) Applicant (for all designated States except US): ITALTEL SOCIETA' ITALIANA TELECOMUNICAZIONI S.P.A. [IT/IT]; Piazzale Zavattari, 12, I-20149 Milan (IT).			
(72) Inventor; and (75) Inventor/Applicant (for US only): COLLIVIGNARELLI, Marco [IT/IT]; Via Verdi, 2, I-28040 Dormelletto (IT).			
(74) Agent: GIUSTINI, Delio; Italtel Società Italiana Telecomunicazioni S.p.A., P.O. Box 10, I-20019 Settimo Milanese (IT).			

## (54) Title: METHOD AND UNIT FOR THE RECOMPOSITION OF ATM CELL FLOWS



## (57) Abstract

Method and unit for the recomposition in a single flow of two or more redundant cell flows outputted from redundant switching fabrics (ASF-a, ASF-b) of an ATM telecommunication system. The method includes the steps of labelling each cell being inputted into the switching fabrics (ASF-a, ASF-b) with a time label (TS) and of storing the cells of the output of the switching fabrics (ASF-a, ASF-b) in a cell buffer (CB). The cells are extracted from said cell buffer when a time span is elapsed, starting from the moment of input of said cell into said fabric, which is equal to a maximum suitable amount of time (MD). The recomposition unit comprises said cell buffer (CB) with optimized capacity as a function of the maximum number of cells which can really be present at the same time in the unit itself.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## Method and unit for the recomposition of ATM cell flows

5

Field of the Invention

This invention relates to a method and a unit for the recomposition in a single flow of two or more redundant cell flows outputted from redundant switching fabrics (ASF-a, ASF-b) of an ATM telecommunication system.

The ATM communication systems use a particular digital technique (ATM: Asynchronous Transfer Mode) for the transmission and/ or the switching of voice signals, video signals and data.

The ATM technique has been identified as the most promising solution for the multiplexing and switching in the evolution of public communication networks towards the future broadband integrated services digital network (B-ISDN) as this technique allows a very large flexibility in comparison with former systems.

On an ATM link a certain number of connections may be simultaneously active, called virtual channels (VC) and virtual paths (VP). The virtual channel is a logical association between the end points of a link that enables the unidirectional transfer of cells over that link, while the virtual path is a logical association of virtual channels between the end points of a link that enables the unidirectional transfer of cells belonging to a set of virtual channels to take place over that link.

In the ATM technique the informations concerning the different subscribers and services are organized as temporary contiguous units with a fixed length of about 400 bits called cells. The bits of the cells are subdivided (figure 1a) in a cell payload (CP) containing the informations to be transported from one subscriber to the other, and in a cell header (CH) containing the

informations necessary for the routing of the cells through the switching network and other service informations.

A switching system in ATM technique consists in particular (figure 2) of a switching fabric ASF and of a certain  
5 number of exchange terminations (ET) the ATM links are connected to.

The ATM cell is received by the inlet ATM link and at the input exchange termination (ET) it is submitted to a format conversion. The new proprietary cell format (figure 1b)  
10 includes fields that allow the routing inside the switching fabric ASF and, in a more general way, the control inside the system. The switching fabric is realized as a multistage and multiple path network, through which the cell is sent to the output exchange termination (ET) of  
15 destination, according to the data contained in the proprietary header.

AT the output exchange termination (ET) the conversion from the proprietary format to the ATM standard format is carried out, and then the cell is sent to the outlet ATM  
20 link.

In order to improve the reliability of the system a redundancy pattern is used providing for the use of two ATM switching fabrics (ASF-a, ASF-b) being contemporaneously operative and transited by said cell flows.

25 If no failure is present, the performance will be improved regarding the cell loss rate as well as the bit error rate; in fact it might be observed that a lost cell or a cell being somehow deteriorated inside a switching fabric may (with a good deal of probability) arrive correctly from the  
30 other switching fabric.

If some failure is present within one of the two switching fabrics, the redundancy pattern arrangement makes it possible to overcome the transitory situation of finding, isolating and/or substituting the failed part without any  
35 significant reduction in the system performance (in terms of loss rate, etc.), and thus without having to declare one

or more active connections out of service. Moreover the applied redundancy pattern arrangement makes it possible to guarantee correct performance even if there are contemporaneously failures on both switching fabrics as far as they do not concern the same elements.

5 The possibility to release the system performance from the intervention rapidity of the error identification and reconfiguration mechanisms as well as to make less critical the technical times required for the substitution of the  
10 damaged parts is very important in this type of systems, due to the high transmission rates and the very high number of connections which could be involved contemporaneously by even only one single failure, especially in applications providing for a random cell distribution mechanism inside  
15 the switching fabrics.

At the input of the switching fabrics (figure 3) there is a recombination unit (DPRC: Duplicated Paths Re-Combiner) realizing for each cell of the single flow present at the proper input two identical copies and sending one of them  
20 to a switching fabric (ASF-a) and the other one to the other switching fabric (ASF-b) via two separate outputs. Inside the switching fabrics the cell copies may meet with different transit delays, as each stage of the ASFs introduces a variable delay and also because, in certain  
25 applications (using a routing with random distribution), the path followed by a copy of the cell inside a fabric may be different from that one followed by the other copy of the same cell inside the other fabric. Therefore the two copies of the same cell being inputted into the two fabrics  
30 at the same time may arrive at the output of the fabrics at different times.

At the output of the fabrics there is a unit DPRCo connected through an input link (DILa) to a fabric (ASF-a) and through the other input link (DILb) to the other fabric  
35 (ASF-b) performing the selection mechanism between the

cells coming from both fabrics, thus supplying to the output a unique cell flow.

#### Background Art

5 Different selection mechanisms are known which are used in the redundancy arrangements of telecommunication equipments in order to exclude the part which nomore guarantees the correct operation and to take the informative flow of interest from the redundant part which was kept as a  
10 reserve until that time. Generally, we talk about stand-by approach and hot stand-by approach according to the fact that the redundant part of the reserve is excluded or not during the normal operation of the part powered-on i.e. in operation. With the stand-by approach it is generally  
15 necessary to accept a longer intervention phase than with hot stand-by approach, and all the informative flow interesting the part to be excluded may be lost unless there are interventions at higher level. On the other hand with the hot stand-by approach only the informative flow  
20 during the switching phase from the damaged part to the reserve part is lost. However, in both cases the system performances are degraded (in terms of services provided to the subscriber), since an intervention by the controller of the selection mechanism will be necessary to diagnose and  
25 correct the failure situation.

In the case of ATM system, the usual times for diagnosis and intervention must be considered unacceptable due to the high transmission rates of the informative flows within the switching system (equal to 360 Mbit/s and more). Therefore  
30 it is necessary to use a switching fabric redundancy arrangement based on two identical structures, operating simultaneously on the same cell flows, both in operation. At the output of the switching node, the selection between the two switching fabrics is carried out cell by cell. In  
35 this case it is possible to talk about an active hot stand-

by redundancy approach, wherein one switching fabric is cell by cell the reserve of the other one.

The European patent application Nr. 91303027.6 discloses a method according to which a sequence number is assigned to the cell valid independently for each connection.

5 At the output exchange terminations (ET), the selection unit controls if it has already received a cell from the same connection with the same sequence number or a lower sequence number in which case the cell is rejected; if, on  
10 the other hand, the cell has a sequence number one greater than that of the cell received before the unit accepts it and transmits it to the output link as soon as possible. In the selection mechanism also an identification code of the switching fabric, from where the accepted cell arrives,  
15 intervenes.

The main drawback of the above described mechanism consists in the fact that it requires a decision about the acceptance of the cell at the moment at which it is received at the input of the selection unit. This also  
20 happens if the received cell is not the one following immediately the one received before (or if it turns out to be damaged). Therefore once a cell is accepted, the preceding cells or those with lower sequence numbers are put aside without considering the possibility that the  
25 missing cells may arrive later from the other switching fabric.

Another drawback of said prior art method is that it requires that the cells of each VC/VP arrive at the output of the switching fabric by the same sequence with which  
30 they entered it.

#### Objects of the Invention

It is an object of this invention to realize a method and a unit for the recomposition in one single flow of two or  
35 more redundant cell flows at the output of at least two redundant switching fabrics (ASF-a, ASF-b) of an ATM

telecommunication system suitable to carry out this operation without losing other cells if not those turning out contemporaneously lost inside all redundant switching fabrics.

- 5 It is a further object of this invention to realize a method and a unit suitable to carry out this operation efficiently in terms of use of the required memory and of complexity of the control concerning the management of this memory.
- 10 Specifically, whenever the ATM telecommunication system requires operations for the reconstruction of the original cell flow sequency order, it is the object of this invention to provide a unit for the recomposition of several redundant flows suitable to carry out efficiently
- 15 these further functions without useless circuit duplication. In fact, using the method this invention is based on, there is no necessity - neither upstream nor downstream - to use a unit for the reconstruction of the original cell flow sequence order, for example of the type
- 20 described in PCT/EP92/02565. So the recombination method according to the invention does not require that the original sequence is respected and it is moreover suitable to restore inherently said sequence (for each VC/VP) in case that this has been altered while passing through the
- 25 switching fabric.

#### Disclosure of the Invention

- These and other objects are achieved by the invention consisting of a method for the recomposition into a unique
- 30 flow of two or more redundant cell flows being outputted from redundant switching fabrics of an ATM telecommunication system, characterized in that it includes the steps of:

- 35       labelling each cell entering the switching fabric with a time label;



- defining for each cell at the output of the switching fabric the time spent inside said switching fabric;
  - storing in a cell buffer each cell of the cells of the output of the switching fabric which
    - 5 i. does not turn out to be the copy of a cell already stored in said cell buffer, or
    - ii. turns out to be the copy of a cell already received but with wrong payload,and for which time span has passed, starting from the moment of input of said cell into said fabric, which is10 less than a maximum suitable time amount;
  - extracting and deleting from said cell buffer, and transmitting to the output channel, each cell, stored in said cell buffer, having passed time span from the15 moment of input of said cell into said fabrics which is equal to said maximum suitable time amount.
- Moreover the invention relates to a unit for the recomposition in a unique flow of two or more redundant cell flows being outputted from redundant switching fabrics20 of an ATM telecommunication system, said unit being located in an exchange termination between the ATM links connecting to said redundant switching fabrics and a cell header processor, wherein said unit comprises an input section receiving the cells from said cell header processor and25 supplying them to said redundant switching fabrics, and an output section receiving the cells coming from said redundant switching fabrics and supplying them to said cell header processor,
- said input section being characterized in that it includes:
- 30 - a labelling unit connected to the cell input channel inserting into the header of each cell an integer number generated by a sequential generator; and
  - a cell copying unit supplying onto the channels connecting to said switching fabrics two or more cell35 flows being identical to the unique cell flow supplied by said labelling unit,

and said output section being characterized in that it includes:

- 5       - an exclusion unit receiving all cells coming from the output channels of said switching fabrics and, for each of said cells, calculating the transit inside delay the switching fabric as the difference between the value of said integer number stored in the header of said cell and the value presently supplied by said sequential generator, and rejecting the cells with a  
10       transit delay greater than the value stored in a register;
- a CAM memory containing updated informations about the presence of the cells stored in said cell buffer;
- a selection unit connected to said CAM memory  
15       supplying, for each cell, a logic value indicating the presence of a copy of such a cell in said cell buffer;
- a calculating unit supplying a value indicating the residual time during which a cell must be stored in said cell buffer;
- 20       - a control unit allowing to insert into said cell buffer each cell for which said logic value supplied by said selection unit indicates the absence in said buffer of a cell copy of such a cell, and suitable for extracting toward the output channel from said cell  
25       buffer each cell for which said residual time value turns out to be zero.

Further advantageous features result from the appended claims.

30       Brief Description of the Drawings

The invention will now be described with reference to a preferred but not restrictive embodiment illustrated by the accompanying drawings where:

- 35       - the Figures 1a and 1b, already mentioned, show the ATM cell valid outside of the system and the proprietary

cell valid on the other side inside of it, respectively;

- Figure 2 , already mentioned, shows schematically the ATM switching fabric duplicated for redundancy, and the exchange terminations connected to it;
- Figure 3 illustrates schematically the structure of an exchange termination ET with a recombination unit in a unique cell flow the two redundant flows;
- Figure 4 illustrates the more detailed block diagram of the recombination unit according to the invention;
- Figure 5 shows a data structure used in an embodiment of the invention to carry out an equalization operation to the maximum delay.

#### Detailed Description of the Preferred Embodiment

It must be noticed that the various connections shown in these figures as a single wire may in fact each one include a plurality of wires.

With reference to Figure 2, an ATM switching node includes a switching fabric ASF duplicated for redundancy purposes (ASF-a and ASF-b) to which different exchange terminations ET1, ET2, ... ETn are connected.

A switching node of this type may be the one described in the patent application Nr. PCT/EP93/01675 filed on June 29, 1993 on behalf of the same applicant entitled "Local and/or transit exchange for a broadband communication network".

The structure of an ET is illustrated in a more detailed way in Figure 3. This includes a CHP block subdivided in an input part CHPi connected to the input link EIL and an output part CHPo connected to the output link EOL.

For reliability reasons the ASF is organized according to a redundancy arrangement with two identical ASFs, ASF-a and ASF-b, operating simultaneously and passed through by the same flows obtained by duplication of the unique flow arriving from the CHP.

In the exchange termination ET and located between the CHP block and the two ASFs, there is a recombination unit DPRC (Duplicated Path Re-Combiner), subdivided in an input part DPRCi and an output part DPRCo. The exchange termination ET  
5 of the switching node operates in this way.

The exchange terminations ET1, ET2, ... ETn receive cell flows entering IN/F1, IN/F2, ... IN/Fn and transmit output cell flows OUT/F1, OUT/F2, ... OUT/Fn. These input and output flows may transit through ATM links, for example  
10 with a capacity of about 150 Mbit/s.

For each exchange termination ET the input part CHPi of the CHP block receives the ATM cell flow from the input link EIL, processes the cells at the ATM level and carries out the format conversion from the ATM one to the proprietary one which is valid inside the system. Dually the output  
15 part CHPo carries out the cell conversion from the inherent node proprietary format to the ATM standard format before sending the cells towards the output link EOL.

The recombination unit DPRC controls the redundant flows sent to the two ASFs and arriving therefrom.  
20

The block diagram corresponding to a preferred embodiment of the recombination unit DPRC is illustrated in figure 4:

- the blocks IN, TSI, FG, CC, OUTa and OUTb and the registers CR1 and CR2 are related to the DPRCi part of  
25 figure 3;
- the blocks INa, FVa, INb, FVb, CS, MT, WTC, DEC, CB, OUT and TBG and the register MDR are related to the DPRCo part of figure 3;
- the block TSG also present in the unit DPRC is shared  
30 by the two parts DPRCi and DPRCo.

The operation of the recombination unit DPRC may be explained as follows.

In the input ET the DPRC unit writes in the cell header of the proprietary cell format the value TS of the time at  
35 which the cell is sent towards the ASF, and realizes the duplication of the cell flow arriving from the CHP adding

at the tail of the cell a FCS code for the correctness control of the payload of the cell itself.

At the output ET the DPRC unit receives the cells arriving from the two ASFs and decides if:

- 5     -     it will accept the cell as not yet received or as the copy of an already received cell with wrong payload CP;

or on the other hand

- 10    -     it will reject the cell as the copy of a cell already received with correct payload CP or because it has not arrived within the acceptable time (transit delay greater than MD).

The cell overcoming the selection step will then be stored in the cell buffer CB for the residual waiting time WT resulting from the comparison between MD and the actual  
15    time the cell spent inside the ASF. Once the cell has exhausted its waiting time in the DPRC unit it will be taken from the cell buffer and sent to the output channel DOL. During the cell extraction step from the cell buffer  
20    CB the unavoidable contentions inherent in the maximum delay equalization algorithm are handled.

The recombination method of the two redundant flows and the functions of the single blocks of the DPRC unit according to the invention will be herein described in a more  
25    detailed way, still with reference to Figure 4.

The cell flow arriving from the CHP is received by the input block IN connected to the input channel DIL; then the TSI block inserts the value TS in the header CIB, in the location shown in Figure 1b, being this value generated  
30    in the TSG block. The value TS is represented with a number N of bits so there may be  $2^N - 1$  different values of TS. N is defined in a way that  $2^N - 1$  is greater than the double of the maximum variation MJD of the transit delay across the ASF, in order to be able to identify exactly the two copies  
35    of the same cell at the output of the ASFs; if MJD is evaluated as a function of the 10<sup>-10</sup>- quantile of MD, it is

possible to maintain said dimensioning of N while accepting a given probability of wrong identification of the cell copy, or N may be dimensioned as a function of the absolute value of the maximum transit delay.

- 5 Downstream the TSI block the FG block generates the code FCS for the control of the correct transmission of the cell payload CP through the ASF and adds it to the tail of said cell. Said code is a Cyclic Redundancy Code (CRC) which check will be carried out in the recombination unit DPRC of
- 10 the output ET in order to identify single or multiple errors. The code control FCS makes it possible - at the output of the ASFs - to choose between the two copies of the cell the one which is without errors in the CP field and to keep the transmission performance of the ASFs under
- 15 permanent control; if on the other hand one of the two copies is received with errors in the CP field and the other one is not received, or if both copies have errors in the CP fields one copy of the cell is anyhow forwarded to the CHP as it is not allowed that a cell can be rejected
- 20 because of error in its payload (in fact this is in charge of the ATM Adaptation level: see the recommendation CCITT I.363 - "ATM adaptation layer (AAL) specification").
- After the block FG the cell passes to the copy block CC realizing the duplication of the cell flow in order to
- 25 obtain two identical redundant flows being separately transmitted to ASF-a and ASF-b through the output blocks OUTa and OUTb connected to the output channels DOLa and DOLb. The duplication function may be disabled either globally or on a cell basis in order to exclude completely
- 30 or partially one of the two ASFs depending on the content of the registers CR1 and CR2. In particular, the register CR1 is used to exclude completely one of the two ASFs depending on a code contained therein, which writing is the task of the ET controller; on the other hand the register
- 35 CR2 is used to exclude one of the two ASFs on a cell basis

depending on a code DC contained in the cell header CIH (figure 1b) and then stored in the register CR2.

In the output ET the DPRC unit receives the cells arriving from ASF-a and from ASF-b through the input blocks INa and INb connected to the input channels DILa and DILb and verifies the correctness of the cell in general and of the CP field in particular (in the blocks FVa and FVb).

The further control of the cells in the DPRC unit is organized according to an internal time base: the block TBG generates for each cell time the pulses P1, P2, ... P7, so as to bracket the different steps which the cell time is subdivided inside the DPRC unit as summarized in the following table:

- from P1 to P2: first cell acceptance examination;
- from P2 to P3: second cell acceptance examination and waiting time WT calculation for the first cell;
- from P3 to P4: waiting time WT calculation for the second cell and first cell writing into the cell buffer CB;
- from P4 to P5: second cell writing into the cell buffer CB and content updating of the table MT for the first cell;
- from P5 to P6: cell extraction from the cell buffer CB and content updating of the table MT for the second cell;
- from P6 to P7: content updating of the table MT for the cell extracted from the cell buffer CB.

If there is only one cell accepted or none in the cell time under consideration, or if no cell is extracted from the cell buffer CB the corresponding operations, provided in the aforementioned steps, are not carried out.

The cell acceptance examination is carried out in the selection block CS. First of all there is the control if the cell underwent any transit delays greater than the suitable maximum value MD, and the cell will be rejected if the transit delay is greater than said value; said control

is carried out in cooperation with the delay control block MDC connected to the register MDR containing the suitable maximum delay value. The content of this register may be cabled in HW or preferably may be programmed by the ET controller.

5 Once this phase is finished, the selection block CS verifies that none of the two cell copies have already been received, one of which therefore would already be stored in the DPRC unit; in the affirmative case the cell is  
10 accepted. If, on the other hand, a copy of the cell has already been stored in the cell buffer CB, in order to be able to decide whether to accept or to reject the second copy of the cell, the selection block controls if the last one arrives from the same ASF (in this case the cell is  
15 rejected and a duplication error is notified), and if the CP field of the copy received before turns out to be wrong, then the new cell is accepted and replaces the copy already present in the cell buffer CB.

In order to identify the two copies of the same cell, the  
20 block CS examines the identifier OCI of the connection which the cell belongs to, and the TS value, both contained in the proprietary header (figure 1b) being in fact the same for the two copies of the same cell.

The control of the presence of a cell copy in the unit DPRC  
25 is carried out by means of a matching table MT where the OCIs and TSs of the cells stored in the cell buffer CB are stored; the fields linked to each OCI and TS pair contain the following indications:

- the ASF from which the cell comes RP,
- 30 - the result EP of the correctness verification of the field CP,
- the address CBA of the location in the cell buffer CB where the cell is stored.

Said table can be implemented by a CAM (Content Adressable  
35 Memory), which can be accessed by supplying part of the content (OCI plus TS) obtaining as a reply the indication



MF (Match Flag) about the presence of said content and the fields RP, EP and CBA associated to it.

At this point the cell control differs according to the fact that the cell is or not already present in the DPRC unit.

5 If the accepted cell is a new one (MF not asserted), the WTC block evaluates the extent of the residual time WT, for which the cell must be stored in the cell buffer CB in order to make the arrival of both copies of the same cell possible, while accepting a sufficiently low probability  
10 (for example equal to  $10^{-10}$ ) that no cell arrives. Said value WT is calculated as the difference (module N) between the maximum delay MD and the age of the cell, this one being obtained by comparing the value TS contained in the cell and the value T supplied by the block TSG. The  
15 residual waiting time WT value is therefore supplied to the maximum delay equalization control block DEC which controls the cell insertion in the memory and programs cell extraction once their storing time is over, controlling in  
20 a suitable way the unavoidable contentions inherent in the maximum delay equalization algorithm. The block DEC provides the address A of a free location in the cell buffer CB, and updates the queues of the cells present in said memory, as illustrated in details below.

25 If the cell is not a new one (MF already asserted) but it has been accepted in substitution of a cell already present in the cell buffer CB with wrong CP field, the block WTC is not activated to define the extent of the remaining time WT, as the cell simply replaces the one already stored in  
30 the cell buffer CB at the address CBA for which the moment of extraction has already been scheduled. Therefore the time WT will not be supplied to the block DEC, but on the contrary the value CBA, so that the operation of finding a new free location in the cell buffer CB and updating the  
35 queues of the cells present in said memory is not carried out.

The updating of the content of the table MT is performed by supplying the values OCI and TS of the accepted cell and the indications concerning the ASF the cell comes from, the correctness of the field CP, and the address of the cell  
5 buffer CBA where it has been stored, and they will be memorized respectively in the fields RP, EP and CBA linked to said OCIs and TSs.

Every cell time, the block DEC evaluates if there is at least one cell in the cell buffer CB for which the storage  
10 time is finished; if this happens, the block DEC indicates the address A of the cell to be extracted from the cell buffer CB, then the cell will be transmitted to the output channel DOL via the output block OUT.

The address A is given to the table MT so that the content  
15 will be updated in order to consider the new situation of the cell buffer CB.

A particular situation may occur when at the inputs INa and INb of the unit DPRC the two copies of a same cell turn out to be present simultaneously, thus identified by the same  
20 OCI and TS values. In this case the selection block CS, after having controlled the transit delay, carries out a research on the table MT in order to control if, due to a malfunction, one of the two copies is duplicated (in this case this one will be rejected and reported as an error).  
25 Since, during normal operation there will be no cell with the same OCI and TS in the cell buffer CB, the block CS will carry out the selection between the two cell copies based on the correctness of their CP field. If both cells have a correct CP field, or both are wrong, the CS block  
30 selects casually one of the two and rejects the other one. The operations involving the following blocks are carried out for the accepted cell as described before.

Referring to figure 5, the internal organization and the fashion of operation of the DEC block are here described in  
35 more detail concerning the cell insertion into the cell buffer CB and their extraction depending on the time spent

in said memory in order to realize the equalization to a maximum predefined value of the cell transit delay, passed from the moment of their transmission by the DPRC unit of the input ET to that one of their transmission by the DPRC unit of the output ET.

The organization and the dimensioning of the data structure of figure 5 are essentially influenced by the value MJD of the difference between the  $10^{-10}$  - quantile of the maximum transit delay (MD) and the minimum transit delay (Dm) of the fabrics ASFs:

$$\text{MJD} = (\text{MD} - \text{Dm}) \quad (1)$$

expressed in number of cell times  $T_c$  or time-slots.

The transmission speed of the input channels DILA and DILb and the output channels DOL are assumed to be equal, therefore in one time-slot two cells may be stored in the cell buffer CB and one cell may be extracted from said memory if there is a cell having passed a time equal to MD since its entrance in the fabric ASF even if on the average only one cell is accepted by the selection mechanism placed at the input of the DPRCo part of the DPRC unit. Apart from the unavoidable contentions a cell makes a position free in the buffer at the maximum after (MD-Dm) time-slot from its entrance. Where all cells meet with a delay Dm in the fabric ASF (slightly loaded fabrics), in the cell buffer CB the maximum number of cells is stored up, which is equal to:

$$(\text{MD} - \text{Dm}) \quad (\text{cells}) \quad (2)$$

independently from the ASF they come from. In fact in (MD - Dm) time-slot  $2 * (\text{MD} - \text{Dm})$  cells may be received from the two ASFs, but only (MD-Dm) cells can be stored in the cell buffer CB, as only one of the two copies of the same cell is stored in said buffer.

The predisposition, or scheduling, of the cell transmission the output from the cell buffer CB is realized by controlling the moments of insertion IS and extraction OS via a list addressed by pointers (figure 5).

A first pointer, called PZOS (Pointer Zero Output Slot), is used to control the global time, therefore it is incremented every cell time  $T_c$ . This pointer identifies both the input slot IS where the cell is received and, at the same time, the output slot OS. This pointer controls moreover the output slot OS for which the transmission of the cells is scheduled, cells which must nomore wait (contentions apart) before being transmitted on the output channel either because they have already spent MD cell times in the fabric (therefore they have a residual waiting time WT equal to zero), or because they have exhausted the waiting period in the cell buffer CB.

A second pointer, called PSOS (Pointer Scheduled Output Slot), is used during the receiving phase and it identifies the OS the transmission of the cell actually being received is scheduled for. PSOS identifies the queue in which the cell being received will be inserted. PSOS may coincide with PZOS: this happens for the cells having spent inside the fabric a time equal to the maximum allowable time (therefore WT results equal to zero).

If there were no contentions of the OS, the two pointers PZOS and PSOS would be sufficient for the controlling of the cell reception and transmission events in the cell buffer CB; PZOB would identify the queue (or better the cell) served in transmission and PSOS the queue served in reception. For controlling of the contentions there is a third pointer, called PVOS (pointer Virtual Output Slot), identifying the output slot OS being "virtually" occupied by all cells disputing it, if their contemporaneous transmission during the contested OS were possible. This pointer is used during the transmission phase to address the queue used in transmission, therefore it is only incremented when this queue is empty, in order to address the next queue containing the cell to be transmitted; if there are no cells to be transmitted in the following OS the PVOS is set equal to the PZOS.

Due to the limits set to the dimensions of the pointers the input slots IS and the output ones OS can only be addressed within a time-window subdivided in time-slots of amplitude  $T_c$  which must be large enough to allow the transmission scheduling of the cells being  $(MD - Dm)$  time-slot away from PZOS, without obtaining a superseding among the cells in the cell buffer CB. At the maximum a queue may contain  $(MD - Dm)$  cells: this happens when all cells present in the cell buffer have entered the ASF fabric at the same time. Therefore PZOS and PSOS may differ by  $(MD - Dm)$  positions, and so the number of addressable OS and, as a consequence, the dimension of the window must be greater or equal to:

$$2 \cdot (MD - Dm) + 1 \quad (\text{time-slot}) \quad (3)$$

The control of the time-slot is implemented by the TSlotList memory shown in figure 5, addressed by the pointers PZOS, PSOS and PVOS, which represents the portion of time axis within which it is possible to address the slots. Each record of this list, that is each location of the memory making up the list, is therefore assimilable to a time-slot, and so the record number of the TSlotList is defined by the Eq.3.

To resolve the contention situations of the OS, the disputing cells are organized in queues, therefore it is possible to have at the most  $(MD - Dm)$  cells memorized in:

- $(MD - Dm)$  queues of 1 cell;
- one queue of  $(MD - Dm)$  cells;
- a whatever intermediate situation between the two previous extreme situations.

With a conventional approach with separate queues there should be a queue connected to each output slot OS dimensioned on the maximum number of cells this one might be able to contain. The number of OS slots is at least equal to  $2 \cdot (MD - DM) + 1$ , and each queue might contain up to  $(MD - Dm)$  cells, therefore, in order to memorize the cell by this approach, it would be necessary to reserve an area of a size equal to:

$$2*(MD - Dm)^2 + (MD - Dm) \quad (\text{cells}) \quad (4)$$

However, the number of active queues at the same time is equal to the number of OS for which the transmission of at least one cell has been scheduled, so there can be a maximum of  $(MD - Dm)$  queues (see relation 2).

In the block DEC an approach of shared queues is used in order to obtain a considerable saving for the size of the cell buffer CB, whose dimensions equal are to:

$$(MD - Dm) + 1 \quad (\text{cells}) \quad (5)$$

corresponding to the maximum number of cells which can be really present; in this way the dimension of the cell buffer CB is traffic independent, so that the probability of additional loss due to buffer overflow turns out to be zero.

The elements of each queue are organized in stacks; in fact the queues are served according to the rule LIFO discipline (Last In First Out) to simplify the management, and so it is sufficient to know the position of the last cell of the queue and the link to the cell occupying the previous position, while a FIFO management needs informations concerning the first and the last cell in the queue.

To manage the queues, a pointer is stored in each TSlotList record, called PLC (Pointer Last Cell) which addresses the last cell of the queue. An empty queue is identified by a value CAP1 connected to PLC operating as a terminator indicating the end of the queue. In a more general way, the queues of the lists are separated by empty elements constituted by terminators set up as univocally recognizable codes.

The informations necessary for the management of the cell buffer CB are stored in lists, and such a list is connected to each active queue. A linked list architecture is used so that even the memory reserved for the lists, called LListMem (Linked List Memory) is shared by the different queues, and it has the same depth as the cell buffer CB, that is:

(MD - Dm) + 1 (record) (6)

5 This memory is also used to store the location list of the free cells (Free Cell List), the first of which is pointed by FlpHead (Free Cell List Pointer Head) contained in a special register.

For the management of the stacks, a pointer is connected to each element of the queue, called PPC (Pointer Previous Cell), sending the cell in the previous position of the same queue. The first cell of each queue is identified by a value CAP1 connected to PPC operating as a terminator since  
10 this cell is not preceded by other cells.

In order to allow PVOS to address the queue to be served in connection with the following OS once the present queue is exhausted to each active queue included between PVOS and PZOS, a pointer PNQ (Pointer Next Queue) is connected pointing to the next active queue. If there is no active queue included between PZOS and PVOS a part from that one actually addressed by PVOS (i.e. served in transmission), PNQ is equal to PZOS. Through PNQ a structure of linked  
15 lists is thus implemented in which the elements of each list are the queues "validated" by PZOS, where by this term those active queues (i.e. which contain at least one cell) are indicated which the cells having already exhausted the waiting time in the DPRC belong to. Practically, the  
20 "validated" queues are all the active queues already addressed by PZOS and not yet by PVOS. The separation element between the queues has been indicated by CAP2 as this is set up - for addressing reasons - by one bit more compared to the previous separators.

30 In figure 5 the memory block NxtQList (Next Queue List), storing the list of the pointers PNQ to the queues to be served after the present one (i.e. pointed by PVOS), and the pointer PLQ (Pointer Last Queue), addressing the last active queue having been validated by PZOS, are  
35 illustrated. As between PVOS and PZOS there may be a maximum of (MD - Dm) queues, the NxtQList has the same

depth as the LListMem, that is  $(MD - Dm+1)$  records, therefore it is addressed by the pointer PZOS\* being equal to the pointer PZOS divided by 2, that is to the n-1 less significant bits of PZOS, this one being represented by n  
5 bits.

Even if the invention has been described with particular reference to a preferred embodiment and to a preferred way of use, it is not restricted to what illustrated, but it is intended to cover all obvious variations and/or  
10 modifications evident to those skilled in the art.



**CLAIMS**

1. Method for the recomposition in a single flow of two or more redundant cell flows being outputted from redundant switching fabrics (ASF-a, ASF-b) of an ATM telecommunication system, characterized in that it includes the steps of:
- labelling each cell being inputted into the switching fabrics (ASF-a, ASF-b) with a time label (TS);
  - 10 - defining for each cell at the output of said switching fabrics (ASF-a, ASF-b) the time spent inside said switching fabric;
  - storing in a cell buffer (CB) each cell of the cells at the output of said switching fabrics (ASF-a, ASF-b) which:
    - 15 i) does not turn out to be the copy of a cell already stored in said cell buffer (CB), or
    - ii) turns out to be the copy of a cell already received but with wrong payload (CP),
  - 20 and for which a time span has passed, starting from the moment of input of said cell into said fabric which is less than a maximum suitable time amount (MD);
  - extracting and deleting from said cell buffer (CB) and transmitting to the output channel (OC) each cell 25 stored in said cell buffer (CB), having passed a time span, starting from the moment of input of said cell into said fabric, which is equal to said maximum suitable time amount (MD).
- 30 2. Method according to claim 1, characterized in that it includes the steps of:
- adding to each cell entering the switching fabrics (ASF-a, ASF-b) a CRC code (FCS) for controlling the correct transmission of the cell payload (CP);
  - 35 - storing in said cell buffer (CB) each one of the cells being outputted from said switching fabrics

(ASF-a, ASF-b) which does not turn out to be already stored in said cell buffer (CB) with a CRC code (FCS) indicating the correctness of its payload, and having spent a time span from the moment of the input of said cell into said fabric which is less than said maximum suitable time amount (MD).

3. Method according to claim 1, characterized in that said maximum suitable time amount (MD) is equal to the maximum transit delay of the cells inside said redundant switching fabrics (ASF-a, ASF-b).

4. Method according to claim 1, characterized in that it includes the steps of:

- setting up a first list (TslotList) comprising values (PLC) representing the locations in said cell buffer (CB), said list being addressed by three pointers, a first pointer (PZOS) identifying the input time-slot (IS) wherein a cell is received, a second pointer (PSOS) identifying the output time-slot (OS) for which the transmission in output of the received cell is scheduled, and a third pointer or virtual pointer (PVOS) identifying the output time-slot in which more than one cell could contend for the transmission, respectively;
- setting up a second linked list (LListMem), and addressed by values (PLC) contained in said first list (TslotList), in said second list (LListMem) queues being set up with elements (PPC) each one of them addressing the cell in the previous position of the same queue;
- setting up a third list (NxtQList) comprising the pointers (PNQ) pointing to the next queues included between said first (PZOS) and third (PVOS) pointers, and active, i.e., containing at least one not empty element;
- transmitting, in a given output time-slot, the last cell of the eventual queue relative to said time-slot,

and in the following time-slots the other cells of the same queue, with inverted order to that one according to which they have been put into the queue.

5        5. Method according to claim 4, characterized in that the length of said first list (TslotList) is equal to  $2*(MD-Dm) + 1$ , where Dm is the minimum number of cell times necessary for the transit inside said switching fabrics (ASF-a, ASF-b).

10       6. Method according to claim 4, characterized in that the length of said second (LListMem) and third lists (NxtQList) is equal to  $(MD-Dm + 1)$ .

15       7. Method according to claim 4, characterized in that the queues of said second (LListMem) and third list (NxtQList) are separated by empty elements comprising terminators in the form of univocally recognisable codes (CAP1, CAP2).

20       8. Method according to claim 4, characterized by providing a further pointer (PLQ) connected to the third list (NxtQList) addressing the last active queue having been validated by said first pointer (PZOS).

25       9. Unit for the recombination in one single flow of two or more redundant cell flows being outputted from redundant switching fabrics (ASF-a, ASF-b) of an ATM telecommunication system, said unit (DPRC) being located in an exchange termination (ET) between the ATM links (DOLa, DOLb, DILa, DILb), connecting to said redundant switching fabrics (ASF-a, ASF-b), and a cell header processor (CHP), wherein said unit comprises an input section (DPRCi) receiving the cells from said cell header processor and  
30       supplying them to said redundant switching fabrics (ASF-a, ASF-b), and an output section (DPRCo) receiving the cells coming from said redundant switching fabrics (ASF-a, ASF-b) and supplying them to said cell header processor, said input section being characterized by comprising:

35       - a labelling unit (TSI) connected to the cell input channel (DIL) inserting into the header (CIH) of each

cell an integer number (TS) generated by a sequential generator (TSG); and

- a cell copying unit (CC) supplying onto the channels (DOLa, DOLb) connecting to said switching fabrics (ASF-a, ASF-b) two or more cell flows being identical to the unique cell flow supplied by said labelling unit (TSI),

and said output section being characterized by comprising:

- an exclusion unit (MDC) receiving all cells coming from the output channels (DILa, DILb) of said switching fabrics (ASF-a, ASF-b), and for each of said cells, calculating the transit delay inside the switching fabric as the difference between the value of said integer number (TS) stored in the header (CIH) of said cell and the value presently supplied by said sequential generator (TSG), and rejecting the cells with a transit delay greater than the value (MD) stored in a register (MDR);
- a CAM memory (MT) containing updated informations about the presence of the cells stored in said cell buffer (CB);
- a selection unit (CS) connected to said CAM memory (MT) supplying, for each cell, a logic value indicating the presence of a copy of a such cell in said cell buffer (CB);
- a calculating unit (WTC) supplying a value (WT) indicating the residual time during which a cell must be stored in said cell buffer (CB);
- a control unit (DEC) allowing to insert into said cell buffer (CB) each cell for which said logic value supplied by said selection unit (CS) indicates the absence in said buffer (CB) of a cell copy of such a cell, and suitable for extracting toward the output channel (DOL) from said cell buffer (CB) each cell for which said residual time value (WT) turns out to be zero.

10. Unit according to claim 9, characterized in that said input part (DPRCi) comprises a unit (FG) inserting into the header of each cell a cyclic redundancy code CRC (FCS), that said output part (DPRCo) comprises a control unit connecting to each cell stored in said cell buffer (CB) a logic signal indicating the correctness of its payload (CP) according to the value of said code CRC (FCS), and in that said control unit (DEC) allows the insertion in said cell memory (CB) of each cell for which such logic signal supplied by said selection unit (CS) indicates either the absence in said memory (CB) of a cell copy of such a cell or the presence in said memory (CB) of a cell with wrong payload which is the copy of such a cell.

11. Unit according to claim 9, characterized in that said control unit (DEC) comprises:

- a first memory (TslotList) containing pointers (PLC) to the locations of said cell buffer (CB);
- a second memory (LListMem) wherein queues are set up with elements (PPC) each one addressing the cell in the previous position of the same queue;
- a third memory (NxtQList) containing the pointers (PNQ) pointing the queues to be served after the present queue.

12. Unit according to claim 11, characterized in that it includes three registers each of one containing one pointer, a first pointer (PZOS) identifying the input time-slot (IS) wherein a cell is received, a second pointer (PSOS) identifying the output time-slot (OS) for which the transmission to output of the received cell is scheduled, and a third pointer or virtual pointer (PVOS) identifying the output time-slot in which more than one cell could contend for the transmission, respectively.

13. Unit according to claim 11, characterized in that the locations of said second memory (LListMem) are addressed by the pointers (PLC) contained in said first

memory (TslotList) in the same way as what happens for the addressing of said cell buffer (CB).

14. Unit according to claim 11, characterized in that the length of said second (LListMem) and third memory (NxtQList) are equal to  $(MD - Dm + 1)$ , wherein  $(MD - Dm)$  is the maximum floating delay at which the cells may be outputted from the switching fabrics (ASF-a, ASF-b).

15. Unit according to claim 14, characterized in that said maximum floating delay  $(MD - Dm)$  is fixed equal to the  $10^{-10}$ -quantile of the variable component of said delay.

16. Unit according to whatever claim from 11 to 15, characterized in that it includes a further register containing a pointer (FlpHead) indicating the first of the free cell positions in said cell buffer (CB) and the relative location of said second memory (LListMem).

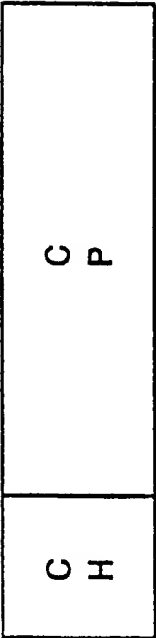


Fig. 1a

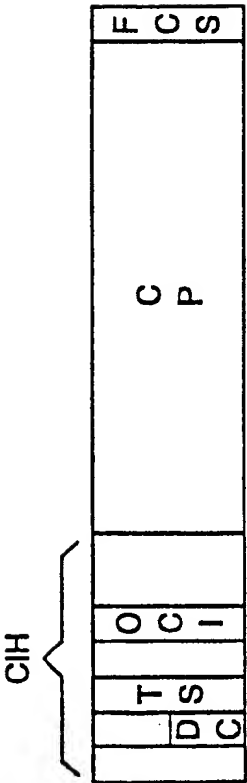


Fig. 1b

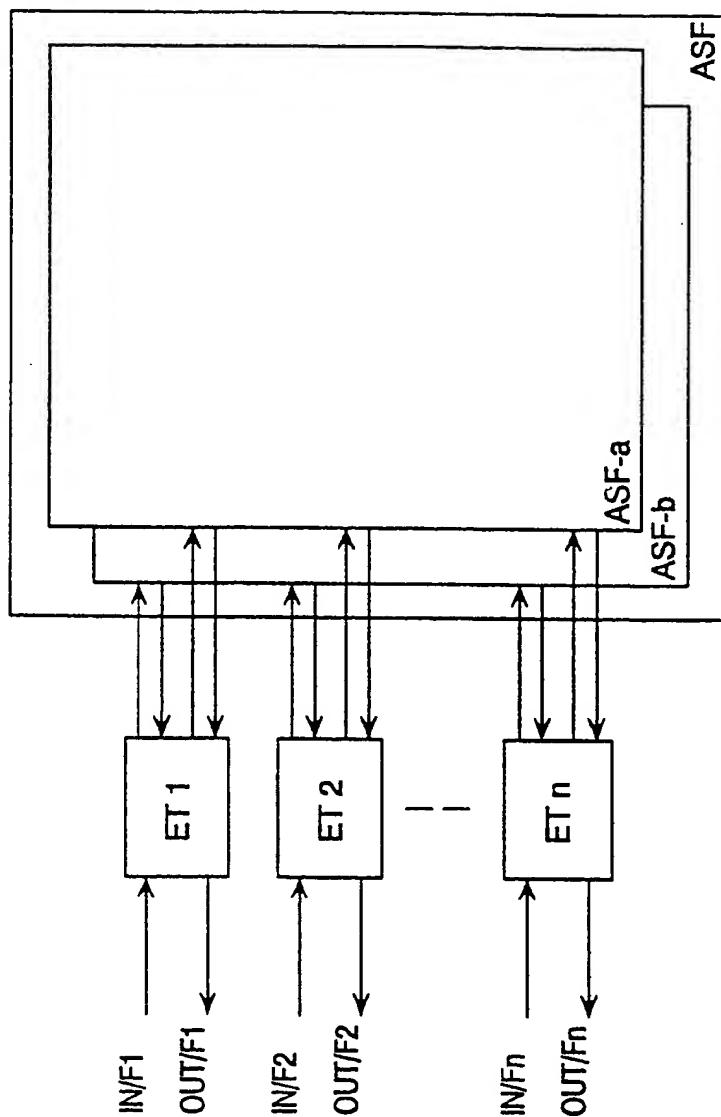


Fig. 2



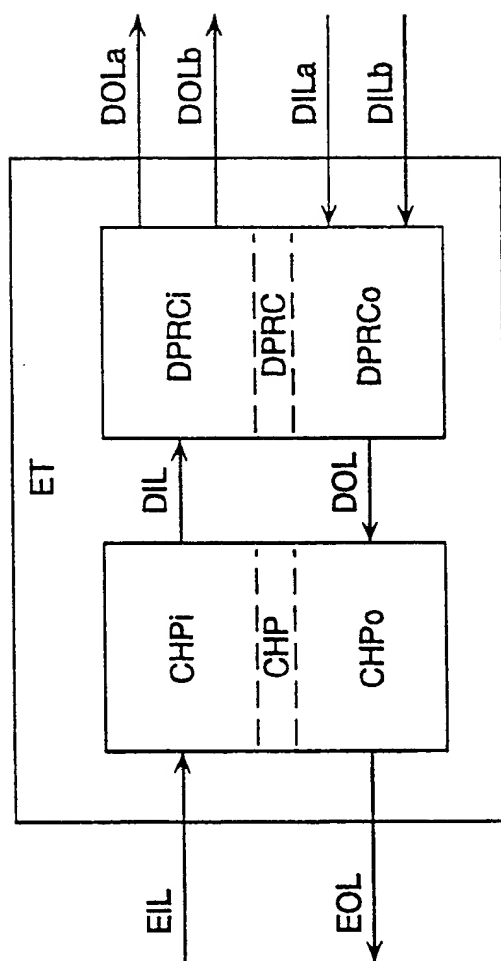


Fig. 3

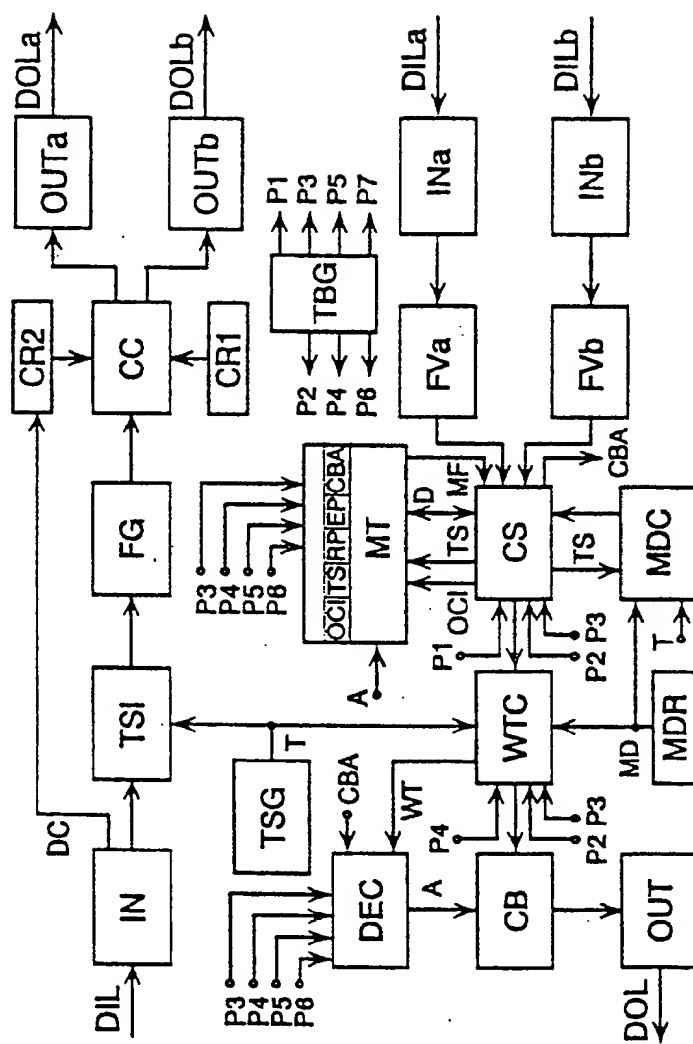


Fig. 4

5/5

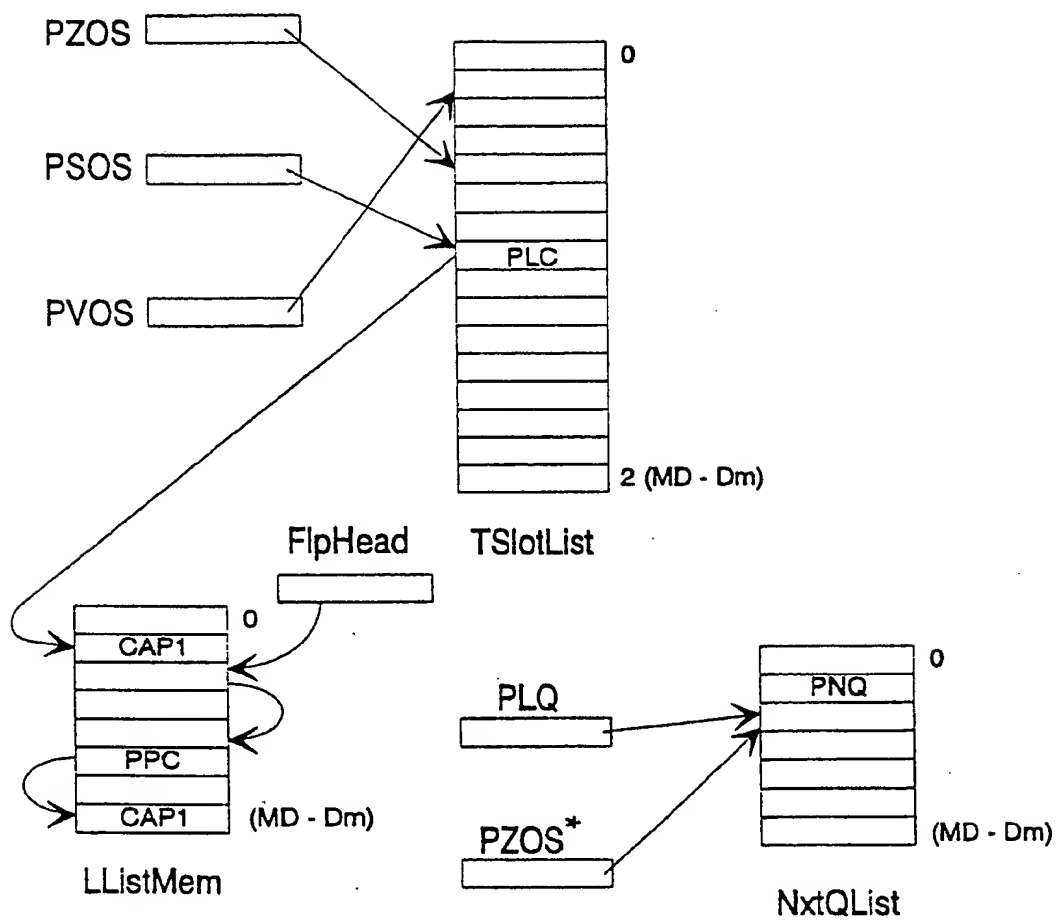


Fig. 5

## INTERNATIONAL SEARCH REPORT

Intern. al Application No

PCT/EP 94/00915

A. CLASSIFICATION OF SUBJECT MATTER  
 IPC 5 H04L12/56 H04Q11/04

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 H04L H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 384 936 (SIEMENS AG) 5 September 1990 see abstract; claims 1-10 ---	1-16
A	INTERNATIONAL SWITCHING SYMPOSIUM 90, vol.5, May 1990, SWEDEN pages 1 - 8, XP130918 M. HERION ET AL see sections 5.1 and 5.2 ---	1-16
A	IEEE INFOCOM 90, June 1990, US pages 105 - 115, XP166288 S. DRAVIDA ET AL see sections ii. and v. -----	1,2,9,10



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\* & \* document member of the same patent family

Date of the actual completion of the international search

6 September 1994

Date of mailing of the international search report

13. 09. 94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax (+ 31-70) 340-3016

Authorized officer

Ali, A

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern. Application No

PCT/EP 94/00915

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0384936	05-09-90	AT-T- 107452	15-07-94
		DE-D- 58907901	21-07-94
		JP-A- 3040627	21-02-91
		US-A- 5325358	28-06-94
-----			

**THIS PAGE BLANK (USPTO)**